6.3 APPLICATION OF THE NEW NWS CLIMATE OUTLOOK IN OPERATIONAL HYDROLOGY

Thomas E. Croley II
Great Lakes Environ.Res. Lab.
Ann Arbor, Michigan

Kenneth E. Kunkel *
Midwest Climate Center
Champaign, Illinois

1. INTRODUCTION

Many forecasters used hydrology models with climatology and the now-discontinued *Monthly and Seasonal Weather Outlook* to make outlooks of basin water supplies. The new *Climate Outlook* of the National Weather Service (NWS) estimates temperature and precipitation probabilities for multiple extended lead times and offers an opportunity to improve water supply forecasts. We developed a system to use the new *Climate Outlook* multiple probabilities to calculate appropriate weighting factors for historical climate scenarios. We describe here the new *Climate Outlook* and its use with hydrology models to make probabilistic outlooks. We derive statistics that use the weights determined from the *Climate Outlook* and we formulate and solve an optimization for finding the weights. We illustrate with an example and discuss the implications.

2. MAKING PROBABILISTIC OUTLOOKS

Beginning with the January 1995 outlook, the NWS Climate Prediction Center provides each month a one-month outlook for the next month and 13 three-month outlooks, going into the future in overlapping fashion in one-month steps. Each outlook estimates probabilities of average air temperature and total precipitation falling within preselected value ranges. The value ranges ("low," "normal," and "high") are defined as the lower, middle, and upper thirds of observations over the period 1961-90 for each variable. The climate outlooks presume that one of only four possibilities exists for the probabilities for each variable: 1) probability of being in the high range exceeds one third and probability of being in the low range is reduced accordingly (it remains at one third for the normal range), referred to as being "above normal," 2) probability of being in the normal range exceeds one third and probabilities of being in the low and high ranges are reduced accordingly and are equal, referred to as being "normal," 3) probability of being in the low range exceeds one third and probability of being in the high range is reduced accordingly (it remains at one third for the normal range), referred to as being "below normal," or 4) skill is insufficient to make a forecast and so probabilities of one-third in each range are used, referred to as "climatological".

* *Corresponding author address:* Kenneth E. Kunkel, Midwest Climate Center, Illinois State Water Survey, 2204 Griffith Drive, Champaign, Illinois 61820-7495.

Users of these climate outlooks can interpret the forecast probabilities in terms of the impacts on themselves through "operational hydrology" approaches. Some operational hydrology approaches consider historical meteorology as possibilities for the future by segmenting the historical record and using each segment with models to simulate a hydrological possibility for the future. Each segment of the historical record then has associated time series of meteorological and hydrological variables, representing a possible "scenario" for the future. The approach then can consider the resulting set of possible future scenarios as a statistical sample and infer probabilities and other parameters associated with both meteorology and hydrology through statistical estimation from this sample (Croley, 1993; Day, 1985). Other operational hydrology approaches use time series models of the historical data to generate the "sample." This increases the precision of the resulting statistical estimates, since large samples can be generated, but not the accuracy. Use of the historical record to directly build a sample for statistical estimation avoids the loss of representation consequent with the use of time series models, but requires a sufficiently large historical record.

The operational hydrology approach uses the tools of statistical sampling as if the set of possible future scenarios were a single "random sample" (i.e., scenarios are independent of each other and equally likely). This means that the relative frequencies of selected events are fixed at values different (generally) than those specified in climate outlooks. Only by restructuring the set of possible future scenarios can we obtain relative frequencies of selected events that match climate outlooks. This restructuring violates the assumption of independent and equally likely scenarios (no random sample) from the point of view of the historical record ("apriori" information). However, the restructured set can be viewed as a random sample ("posterior" information) of scenarios conditioned on climate outlooks. There are many methods for restructuring the set of possible future scenarios (Croley, 1993; Day, 1985; Ingram et al., 1995).

3. BUILDING A STRUCTURED SET

In building an operational hydrology set of possible future scenarios, from which to estimate probabilities and other parameters associated with various meteorological and hydrological variables, consider constructing a structured set that, when treated as a statistical sample, guarantees that probability estimates for certain variables match climate outlooks. That is, we can build a structured set of possible scenarios that gives relative frequencies of average air temperature and total precipitation (over various times in the scenarios) satisfying the apriori settings of the climate outlooks. We can arbitrarily construct a very large structured set of size N by adding (duplicating) each of the available scenarios (in the original set of n possible future scenarios); each scenario numbered i, ($i = 1, ..., n$) is duplicated r_i times. By judiciously choosing these duplication numbers, ($r_1, r_2, ..., r_n$), it is possible to force the relative frequency of any arbitrarily-defined "group" of scenarios in the structured set to any desired value. For example, suppose only five of 50 (10%) twelve-month scenarios beginning in June have an average June air temperature exceeding 30°C, and our apriori setting (from a climate outlook) for this exceedance is 20%. We could repeat each of these five scenarios 9 times and repeat the other 45 scenarios 4 times to build a structured set. This structured set of size 225 (= 5 × 9 + 45 × 4) would then have a relative frequency of 20% of average June air temperature exceeding 30°C (5 × 9 / 225 = 0.2). For sufficiently large N, we can approximate apriori settings at any precision by using integer-valued duplication numbers, r_i. Note also:

$$\sum_{i=1}^{n} r_i = N \qquad (1)$$

By treating the N scenarios in the very large structured set as a statistical sample, we can estimate probabilities and calculate other parameters for all variables. In particular, consider any variable X (either historical meteorological or simulated hydrological); e.g., X might be July-August-September total precipitation, end-of-August soil moisture storage, lake surface temperature on day 55, or average June air temperature. We denote the "event" that a variable X is less than or equal to a value x as $\{X \le x\}$ and the probability of this event as $P[X \le x]$. This probability is estimated, when considering the very large structured set as a statistical sample, by the "relative frequency" of the event in the structured set. The relative frequency of event $\{X \le x\}$ is just the number of scenarios in which the event occurs divided by the set size N:

$$\hat{P}[X \le x] = \sum_{k \in \Omega} \frac{1}{N}, \qquad \Omega \equiv \{k \mid x_k^N \le x\} \qquad (2)$$

where $\hat{P}[\ ]$ denotes a probability estimate, and x_k^N is the value of variable X for the k^{th} scenario in the very large structured set of N scenarios. [Read the set notation in (2) as "Ω is all values of k such that $x_k^N \le x$."] Actually, there are only n different values of X (x_i^n, $i = 1, ..., n$) since these n values were duplicated, each by a number, r_i, to create the N values in the very large structured set. We can rewrite (2) in terms of the original set of possible future scenarios, for any variable X:

$$\hat{P}[X \le x] = \sum_{i \in \Omega} \frac{r_i}{N}, \qquad \Omega \equiv \{i \mid x_i^n \le x\} \qquad (3)$$

Furthermore, we can write other estimators (defined over the large structured set of scenarios as if it was a statistical sample) in terms of the original set. Consider the γ-probability quantile for variable X, ξ_γ; it is defined by:

$$P[X \le \xi_\gamma] = \gamma \qquad (4)$$

The γ-probability quantile, ξ_γ, is estimated, when considering the structured set as a statistical sample, by the m^{th} order statistic, y_m^N, where $m = \gamma N$. Order all values of X in the very large structured set (x_k^N, $k = 1, ..., N$) from smallest to largest to define the order statistics (y_m^N, $m = 1, ..., N$). The probability estimate is then

$$\hat{P}[X \le y_m^N] = \frac{m}{N}, \qquad m = 1, \ldots, N \qquad (5)$$

where $y_m^N = x_{k(m)}^N$ and $k(m)$ is the number of the value in the structured set corresponding to the m^{th} order. [For example, if the third value in the structured set, x_3^N, was the largest ($y_N^N = x_3^N$), then $k(N) = 3$]. Alternatively, (5) can be written as follows.

$$\hat{P}[X \le x_{k(m)}^N] = \sum_{l=1}^{m} \frac{1}{N}, \qquad m = 1, \ldots, N \qquad (6)$$

In terms of order statistics for the original set (y_j^n, $j = 1, \ldots, n$), there are $r_{i(j)}$ identical values of y_j^n in the very large structured set where $i(j)$ is defined similarly to $k(m)$ but for the original set in which $j = 1, \ldots, n$, and $y_j^n = x_{i(j)}^n$. Equations (5) and (6) may be rewritten in terms of the original set of possible future scenarios (for any variable X):

$$\hat{P}[X \le y_j^n] = \hat{P}[X \le x_{i(j)}^n] = \sum_{l=1}^{j} \frac{r_{i(l)}}{N}, \quad j = 1, \ldots, n \qquad (7)$$

Likewise, the mean and variance of variable X over the structured set, $\bar{x}$ and S^2 respectively, become, in terms of the original set:

$$\bar{x} = \frac{1}{N}\sum_{k=1}^{N} x_k^N = \frac{1}{N}\sum_{i=1}^{n} r_i\, x_i^n$$
$$S^2 = \frac{1}{N}\sum_{k=1}^{N}\left(x_k^N - \bar{x}\right)^2 = \frac{1}{N}\sum_{i=1}^{n} r_i\left(x_i^n - \bar{x}\right)^2 \qquad (8)$$

Rewriting (3), (7), and (8),

$$\hat{P}[X \le x] = \frac{1}{n}\sum_{i\in\Omega} w_i, \qquad \Omega \equiv \{i \mid x_i^n \le x\}$$

$$\hat{P}\left[X \le x_{i(j)}^n\right] = \frac{1}{n}\sum_{l=1}^{j} w_{i(l)}, \qquad j = 1, \ldots, n$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} w_i x_i^n \tag{9}$$

$$S^2 = \frac{1}{n}\sum_{i=1}^{n} w_i\left(x_i^n - \bar{x}\right)^2$$

where

$$w_i = \frac{n}{N} r_i \tag{10}$$

Note that

$$\sum_{i=1}^{n} w_i = n \tag{11}$$

and if all $w_i = 1$, then (9) gives contemporary (unstructured) estimates from the original set treated as a statistical sample. Other statistics can be similarly derived.

4. CONSIDERING MULTIPLE OUTLOOKS

Now consider the case of multiple apriori settings (from a climate outlook) to which to match relative frequencies. For example, consider the settings from the new NWS Climate Prediction Center *Climate Outlook*.

$$\begin{aligned}
\hat{P}\left[T_g > \tau_{g,0.667}\right] &= a_g \\
\hat{P}\left[T_g \le \tau_{g,0.333}\right] &= b_g \\
\hat{P}\left[\tau_{g,0.333} < T_g \le \tau_{g,0.667}\right] &= 1 - a_g - b_g \\
\hat{P}\left[Q_g > \theta_{g,0.667}\right] &= c_g \\
\hat{P}\left[Q_g \le \theta_{g,0.333}\right] &= d_g \\
\hat{P}\left[\theta_{g,0.333} < Q_g \le \theta_{g,0.667}\right] &= 1 - c_g - d_g \\
g &= 1, \ldots, 14
\end{aligned} \tag{12}$$

where T_g and Q_g are average air temperature and total precipitation, respectively, over period g ($g = 1$ corresponds to a one-month period, and $g = 2, \ldots, 14$ corresponds to 13 successive overlapping three-month periods), $\tau_{g,\gamma}$ and $\theta_{g,\gamma}$ are, respectively, temperature and precipitation reference γ-probability quantiles for period g, and (a_g, b_g, c_g, and d_g, $g = 1, \ldots, 14$) are the outlook settings. By definition, the reference γ-probability quantiles are estimated from the 1961-90 historical record for each period g. To illustrate (12), consider the June 1995 *Climate Outlook*; there is a one-month June outlook ($g = 1$ or "Jun") and 13 three-month outlooks successively lagged by one month each (g = 2 or "June-July-August" or "JJA," and $g = 3, 4, \ldots, 14$ or "JAS," "ASO," ..., "JJA," respectively). The third and sixth lines in (12) are redundant with the rest of (12) because relative frequencies sum to unity:

$$\begin{aligned}
&\hat{P}\left[T_g \le \tau_{g,0.333}\right] \\
&+ \hat{P}\left[\tau_{g,0.333} < T_g \le \tau_{g,0.667}\right] \\
&+ \hat{P}\left[T_g > \tau_{g,0.667}\right] = 1 \\
&\hat{P}\left[Q_g \le \theta_{g,0.333}\right] \qquad\qquad g = 1, \ldots, 14 \\
&+ \hat{P}\left[\theta_{g,0.333} < Q_g \le \theta_{g,0.667}\right] \\
&+ \hat{P}\left[Q_g > \theta_{g,0.667}\right] = 1
\end{aligned} \tag{13}$$

Since relative frequencies sum to unity, there are four independent settings in (12) for each of the 14 climate outlooks for a total of 56, if all outlooks are used.

Rewriting (12) and (13) in light of the first line of (9),

$$\begin{aligned}
\sum_{i\in A_g} w_i &= a_g n, & A_g &\equiv \{i \mid t_{g,i} > \tau_{g,0.667}\}, & g &= 1, \ldots, 14 \\
\sum_{i\in B_g} w_i &= b_g n, & B_g &\equiv \{i \mid t_{g,i} \le \tau_{g,0.333}\}, & g &= 1, \ldots, 14 \\
\sum_{i\in C_g} w_i &= c_g n, & C_g &\equiv \{i \mid q_{g,i} > \theta_{g,0.667}\}, & g &= 1, \ldots, 14 \\
\sum_{i\in D_g} w_i &= d_g n, & D_g &\equiv \{i \mid q_{g,i} \le \theta_{g,0.333}\}, & g &= 1, \ldots, 14 \\
\sum_{i=1}^{n} w_i &= n
\end{aligned} \tag{14}$$

where $t_{g,i}$ and $q_{g,i}$ are average air temperature and total precipitation, respectively, over period g of scenario i. Alternatively, (14) can be written as follows.

$$\sum_{i=1}^{n} a_{k,i} w_i = e_k, \qquad k = 1, \ldots, 57 \tag{15}$$

where $a_{k,i}$ has the value of 0 or 1 corresponding to the exclusion or inclusion, respectively, of each variable in the above sets, and e_k corresponds to the climate outlook relative frequency settings specified in (12):

$$\begin{aligned}
e_k &= a_k\, n, & k &= 1, \ldots, 14 \\
&= b_{k-14}\, n, & k &= 15, \ldots, 28 \\
&= c_{k-28}\, n, & k &= 29, \ldots, 42 \\
&= d_{k-42}\, n, & k &= 43, \ldots, 56 \\
&= n, & k &= 57
\end{aligned} \tag{16}$$

Ordinarily, all of the Climate Prediction Center climate outlooks may not be used, in which case simply write (15) as

$$\sum_{i=1}^{n} a_{k,i} w_i \;=\; e_k\,, \qquad k = 1, \ldots, m \qquad (17)$$

where $m \le 57$, and the appropriate equations, corresponding to the unused outlooks, are omitted. We must solve (17) simultaneously to find the weights.

Generally, $m \neq n$ and some of the equations may be either redundant or non-intersecting with the rest and must be eliminated. (If $m > n$, then $m - n$ of the equations must be either redundant or non-intersecting. This corresponds to not being able to simultaneously satisfy all climate outlook information with fewer scenarios than there are outlook boundary conditions.) Selection of some for elimination is facilitated by assigning each equation in (17) a priority reflecting its importance to the user. [The highest priority is given to the equation in (17) corresponding to the last line of (14), guaranteeing that all relative frequencies sum to unity.] Each equation, in priority order starting with the next-to-highest priority, is compared to the set of all higher-priority equations and eliminated if it is redundant or does not intersect the set. By starting with the higher priorities, we ensure that each equation is compared with a known valid set of equations, and that we keep higher-priority equations in preference to lower-priority equations. Thus we can always reduce (17) so that $m \le n$. If $m = n$, then (17) can be solved via Gauss-Jordan elimination as a system of linear equations for the weights, w_i, since the equations are now independent and intersecting (in n-space). Else, $m < n$, and (17) consists of the remaining independent intersecting equations.

There are multiple solutions to (17) for $m < n$, and the identification of the "best" set of weights requires the specification of a measure for comparing the solutions. One such measure is the deviation of weights from unity. Solutions of (17) giving smaller values of this measure can be judged "better" than those that do not (and the resulting very large structured set of scenarios is more similar to the original set of scenarios in this sense). We can formulate an optimization problem to minimize this measure in selecting a solution to (17):

$$\begin{aligned} &\min \sum_{i=1}^{n} (w_i - 1)^2 \\ &\text{s.t.} \quad \sum_{i=1}^{n} a_{k,i} w_i \;=\; e_k\,, \qquad k = 1, \ldots, m \end{aligned} \qquad (18)$$

By defining the "Lagrangian" for this problem (Hillier and Lieberman, 1969),

$$L \;=\; \sum_{i=1}^{n} (w_i - 1)^2 \;-\; \sum_{k=1}^{m} \lambda_k \left(\sum_{i=1}^{n} a_{k,i} w_i \;-\; e_k \right) \qquad (19)$$

(where λ_k = the unit penalty of violating the k^{th} constraint in the optimization) and by setting the first derivatives of the Lagrangian with respect to each "variable" to zero,

$$\begin{aligned} \frac{\partial L}{\partial w_i} &= 2(w_i - 1) - \sum_{k=1}^{m} \lambda_k a_{k,i} = 0, \quad i = 1, \ldots, n \\ \frac{\partial L}{\partial \lambda_k} &= -\sum_{i=1}^{n} a_{k,i} w_i \;+\; e_k \;=\; 0, \quad k = 1, \ldots, m \end{aligned} \qquad (20)$$

we have a set of necessary but not sufficient conditions for the problem of (18). Equations (20) are linear and may be solved via the Gauss-Jordan method of elimination. Sufficiency may be checked by inspection of the solution space in the vicinity of the solution.

The solution of (18), may give positive, zero, or negative weights, but only non-negative weights make physical sense and we must further constrain the optimization to non-negative weights. This can be done by introducing non-negativity inequality constraints into (18), (19), and (20). These additional equations would require enumeration of all "zero points" or "roots" of (20) (a root is a solution with zero-valued weights). However, this is impractical since it can involve the inspection of many roots [e.g., for $n = 50$, there are $2^{50} - 1$ roots ($> 10^{15}$)]. Furthermore, non-negativity constraints can result in infeasibility (there is no solution). In this case, additional lowest-priority equations must be eliminated from (17) to allow a non-negative solution. The following two methods, portrayed in Figure 1, provide systematic procedures for finding non-negative weights through elimination of lowest-priority equations. They

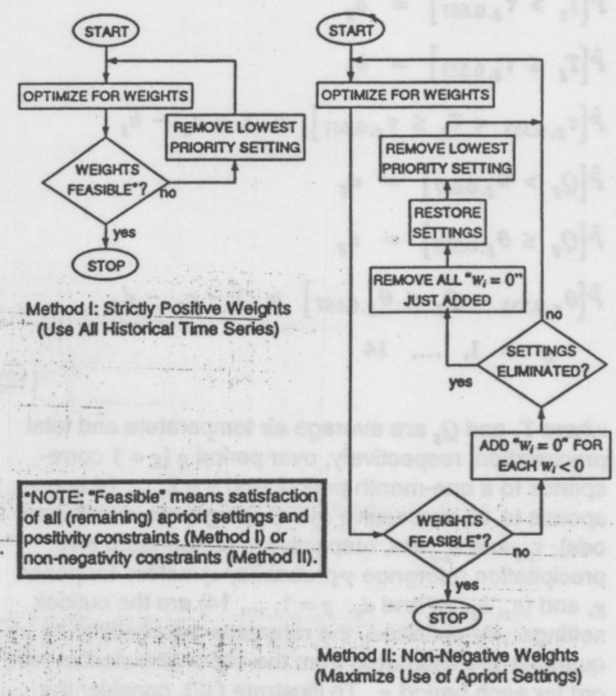

Figure. 1. Reduction/Optimization Algorithms.

also avoid direct use of non-negativity constraints in (18) thus avoiding inspection of the large number of roots that can result. The first method guarantees that only strictly positive weights will result; this means that all possible future scenarios are used (no scenario is weighted by zero and effectively eliminated) in estimating probabilities and other parameters. Alternatively, if we are willing to disallow some of the possible future scenarios (allow zero-valued weights), the second method satisfies more of the apriori settings [more of the equations in (17)] in the event of a negative solution. The reduction of (20) with non-negativity constraints is further described elsewhere (Croley, 1995).

5. MULTIPLE OUTLOOKS EXAMPLE

Consider the following example. GLERL's hydrology models are to be used to estimate the 12-month probabilistic outlook of net basin supply for Lake Superior beginning June 1995 by using the NWS Climate Prediction Center *Climate Outlook* for June 1995. (Net basin supply is the algebraic sum of overlake precipitation, lake evaporation, and basin runoff to the lake.) The outlook will be made by identifying all 12-month meteorological time series that start in June from the available historical record of 1948-93; there are 45 such time series for each meteorological variable. The time series for all meteorological variables will be used in simulations with GLERL's hydrology models and current initial conditions to estimate the 45 associated time series for each hydrological variable. Each set of historical meteorological and associated hydrological time series, corresponding to each segment of the historical record, represent a possible future scenario. The 45 scenarios will be used as a statistical sample in an operational hydrology approach to make the probabilistic outlook. We will incorporate the Climate Prediction Center *Climate Outlook* by using selected period outlook settings as boundary conditions in the determination of weights to apply to our scenario set. We use these weights, through estimates from (9), to make our probabilistic outlook.

We must begin by abstracting historical quantiles of air temperature and precipitation for the Lake Superior basin; these are presented in Table 1 for the periods of interest in making the June outlook. These were estimated from the 1961-90 period in accordance with definitions provided by the Climate Prediction Center for use of their climate outlooks. These quantiles are the basis for interpretation of the Climate Prediction Center's climate outlooks.

The NWS Climate Prediction Center *Climate Outlook* for June 1995 (made 18 May 1995) over the Lake Superior Basin is given in Table 2 in columns two and three. They are interpreted, in accordance with specifications of the Climate Prediction Center [and as described in the section on *Making Probabilistic Outlooks* and in the previous section; see (12)], to construct the

Table 1. Meteorological Quantiles on Lake Superior Basin[a] for Selected Periods.

Period, g	Average Temperature Quantiles $\tau_{g,0.333}$ (°C)	$\tau_{g,0.667}$ (°C)	Total Precipitation Quantiles $\theta_{g,0.333}$ (mm)	$\theta_{g,0.667}$ (mm)
Jun	13.38	14.43	69	106
JJA	15.18	16.29	242	295
JAS	14.49	15.12	240	299
ASO	10.32	11.18	253	282
SON	4.08	5.02	206	247
OND	-3.40	-2.09	178	216
NDJ	-10.30	-9.27	157	190
DJF	-14.19	-12.71	135	151
JFM	-12.68	-10.75	121	135
FMA	-6.86	-4.52	123	146
MAM	0.88	2.13	154	177
AMJ	8.03	8.55	197	230
MJJ	13.04	13.51	234	267

[a]Estimated from 1961-90 daily data over the Lake Superior Basin from 230 meteorological stations averaged spatially.

Table 2. NWS Climate Prediction Center June 1995 *Climate Outlook* Probabilities[a], %.

Period, g	P_T[b]	P_Q[b]	Temperature Probabilities[c] low	*norm.*	*high*	Precipitation Probabilities[c] low	*norm.*	*high*
Jun'95	*0 c*	*0 c*	*33*	*33*	*33*	*33*	*33*	*33*
JJA'95	*0 c*	*0 c*	*33*	*33*	*33*	*33*	*33*	*33*
JAS'95	$\boxed{2\ n}$	0 c	$\boxed{32}$	$\boxed{35}$	$\boxed{32}$	33	33	33
ASO'95	0 c	0 c	33	33	33	33	33	33
SON'95	$\boxed{3\ b}$	0 c	$\boxed{36}$	$\boxed{33}$	$\boxed{30}$	33	33	33
OND'95	0 c	0 c	33	33	33	33	33	33
NDJ'95	0 c	0 c	33	33	33	33	33	33
DJF'95	$\boxed{1\ a}$	0 c	$\boxed{32}$	$\boxed{33}$	$\boxed{34}$	33	33	33
JFM'96	$\boxed{2\ a}$	$\boxed{10\ b}$	$\boxed{31}$	$\boxed{33}$	$\boxed{35}$	$\boxed{43}$	$\boxed{33}$	$\boxed{23}$
FMA'96	$\boxed{1\ a}$	0 c	$\boxed{32}$	$\boxed{33}$	$\boxed{34}$	33	33	33
MAM'96	$\boxed{3\ a}$	0 c	$\boxed{30}$	$\boxed{33}$	$\boxed{36}$	33	33	33
AMJ'96	0 c	0 c	33	33	33	33	33	33
MJJ'96	0 c	0 c	33	33	33	33	33	33
JJA'96	0 c	0 c	33	33	33	33	33	33

[a]For the Lake Superior basin; probabilities expressed as percentages do not appear to sum to unity because of the two-digit round-off used here.
[b]Probability (P_T and P_Q designate temperature and precipitation probabilities, respectively) in excess of 33% in low interval (below normal), in mid interval (normal), or in high interval (above normal); "no forecast" is indicated by "0 c" (climatological).
[c]Probabilities over the Climate Prediction Center's corresponding interval definitions.

Table 3. Boundary Condition Equations (17) for June 1995 Outlook on Lake Superior.

Period, g[a]	k[b]	Interval[c]	Inclusion in interval, $a_{k,i}$, $i = 1, \ldots, 45$[d]	e_k[d]
JAS'95	2	$(\tau_{k,0.667}, \infty)$	110011010001110100100110010000000001000111010	0.32×45
JAS'95	3	$(-\infty, \tau_{k,0.333})$	001100101000001001010001101001010010010000001	0.32×45
SON'95	4	$(\tau_{k,0.667}, \infty)$	100001101010111100001011010001000001100000000	0.30×45
SON'95	5	$(-\infty, \tau_{k,0.333})$	000100000001000001100000101010011000011001010	0.36×45
DJF'95	6	$(\tau_{k,0.667}, \infty)$	100111110101100101001000001000011010001101111	0.34×45
DJF'95	7	$(-\infty, \tau_{k,0.333})$	000000001010011010100001010011100100000000000	0.32×45
JFM'96	8	$(\tau_{k,0.667}, \infty)$	000111000100100100000000100010001011111001111	0.35×45
JFM'96	9	$(-\infty, \tau_{k,0.333})$	010000000010001010000101010001100100000000000	0.31×45
JFM'96	10	$(\theta_{k,0.667}, \infty)$	111011100000000011100011001110100000000110000	0.23×45
JFM'96	11	$(-\infty, \theta_{k,0.333})$	000000011111101000010100000001000011011000111	0.43×45
FMA'96	12	$(\tau_{k,0.667}, \infty)$	000101000100100000000000100010001011111001111	0.34×45
FMA'96	13	$(-\infty, \tau_{k,0.333})$	010000000000000010100101010001100000000010000	0.32×45
MAM'96	14	$(\tau_{k,0.667}, \infty)$	001010100100010000100001000010001000111101111	0.36×45
MAM'96	15	$(-\infty, \tau_{k,0.333})$	010001010001000010100111010000100000000010000	0.30×45
Entire	1		111111111111111111111111111111111111111111111	1.00×45

[a]Period as selected (highlighted) in Table 2.
[b]Period renumbered by priority (1 ≡ highest) as in (17).
[c]Interval as defined in Table 1.
[d]Coefficients in (17) defined for each selected period, k, of the climate outlook, and for each scenario, i, in the historical record.

probabilities associated with the reference quantiles in Table 1; these are given in columns four through nine in Table 2. Highlighted entries in Table 2 denote outlook probabilities designated as significant by the Climate Prediction Center, who suggest that the remainder be estimated from climatology since they have insufficient skill to make outlooks in those cases.

The highlighted entries in Table 2 are used arbitrarily, in priority of their appearance, to make the outlook. These seven outlook settings and the reference quantiles in Table 1 are used with inspection of all 45 scenarios to construct 15 equations represented by (17) in Table 3. Table 4 presents the solution of these equations, found by minimizing the deviation of weights with unity, as in (18). While all 45 scenarios are used (all weights are strictly positive), not all of the selected apriori climate settings can be used. The temperature probability settings for JAS, SON, DJF, and JFM were used while the temperature probability settings for FMA and MAM and the precipitation probability setting for JFM were unused. We could use all seven apriori climate settings if we allowed zero-valued weights. This is done in Table 5 where the scenarios starting in June 1948, 1952, 1953, 1954, 1970, and 1987 are unused. Croley (1995) discusses these alternatives further.

Table 4. Outlook Weights: All Historical Time Series[a].

Year	Weight	Year	Weight	Year	Weight
1948	0.444378	1963	0.259718	1978	1.527387
1949	1.659873	1964	1.527387	1979	1.112034
1950	1.089694	1965	1.112034	1980	1.459070
1951	0.927374	1966	1.183255	1981	1.527387
1952	0.150880	1967	1.089694	1982	0.157130
1953	0.259718	1968	0.982324	1983	1.007623
1954	0.450628	1969	1.659873	1984	1.545569
1955	0.335539	1970	1.192282	1985	1.675279
1956	0.528100	1971	1.104530	1986	1.459070
1957	0.688826	1972	1.675279	1987	0.335539
1958	1.636225	1973	1.098279	1988	1.083444
1959	1.105783	1974	1.112034	1989	0.921124
1960	0.259718	1975	1.621390	1990	0.688826
1961	0.521850	1976	1.536542	1991	0.921124
1962	1.104530	1977	1.104530	1992	0.157130

[a]Solution of (17) with Table 3 values using all historical data years and apriori settings for JAS, SON, DJF, and JFM temperature probabilities; settings for FMA and MAM temperature and JFM precipitation are unused.

Table 5. Outlook Weights: All Apriori Climate Settings[a].

Year	Weight	Year	Weight	Year	Weight
1948	0	1963	0.450000	1978	1.269962
1949	1.060486	1964	1.269962	1979	1.919873
1950	0.312190	1965	0.424136	1980	1.813411
1951	1.008031	1966	1.808557	1981	1.279712
1952	0	1967	1.879379	1982	0.171944
1953	0	1968	1.912046	1983	0.911242
1954	0	1969	2.627675	1984	1.795797
1955	0.357372	1970	0	1985	1.875076
1956	1.137376	1971	0.379306	1986	1.884862
1957	0.977323	1972	1.803624	1987	0
1958	1.355692	1973	1.724416	1988	1.737354
1959	1.264911	1974	0.424136	1989	0.767599
1960	0.025845	1975	1.297178	1990	0.977323
1961	0.825493	1976	0.366735	1991	0.839051
1962	0.460508	1977	2.522282	1992	0.082140

[a]Solution of (17) with Table 3 values using all apriori climate outlook settings highlighted in Table 2.

Finally, as an example for one hydrological variable, the probabilistic outlook for net basin supply (NBS), over the twelve months from June 1995 through May 1996, is given in Table 6. There were 45 values of monthly NBS, corresponding to the 45 scenarios used in the simulation, for each of the twelve months. Each value was multiplied by its respective weight from Table 5, as in (9), to compute various statistics for the probabilistic outlook each month. Selected quantiles from the forecast NBS probability distribution and the mean and standard deviation for each month of the outlook are displayed in Table 6. Since the weights of Table 5 were used, the probabilistic outlook in Table 6 represents use of all selected apriori climate outlook settings.

6. SUMMARY

The operational hydrology approach described herein uses historical information while preserving many of the long-term meteorological probability outlooks provided by NWS's Climate Prediction Center. Other approaches may severely limit the use of historical data to be compatible with climate outlooks or use all historical data only by ignoring these outlooks. The use of a hypothetical very large structured set of scenarios to estimate hydrological outlook probabilities corresponds to the use of the weighted original set of possible future scenarios estimated from the historical record. (Each scenario consists of an actual segment of the historical record and its associated hydrological transformation made with appropriate models.) The building of this hypothetical very large structured set is an arbitrary concept that was useful in defining the weights. The National Weather Service is now considering weighting methods for their Extended Streamflow Prediction operational hydrology approach (Day, 1985) that couple historical time series of precipitation with precipitation forecasts (Ingram et al., 1995).

Still other approaches use time series models, fit to historical data, to generate a large sample, increasing precision but not accuracy in the resulting statistical estimates. Direct use of the historical record to build a sample avoids the loss of representation consequent with time series models. In addition, it may not be clear how to modify time series models to agree with climatic outlooks and still be representative of the underlying behavior originally captured in the time series models. Nevertheless, if time series models are used in building the sample, weighting of this sample, in the manner described herein, to agree with climatic outlooks is straightforward and still could be used.

The determination of these weights involves several choices also made arbitrarily herein. For example, the weights could be determined directly from multiple climate outlooks, as exemplified earlier for a single climate outlook (average June air temperature). This would involve restrictions on the multiple climate outlooks not considered here. The formulation of an optimization problem, used herein, allows for a more general approach in determining these weights in the face of multiple outlooks. However, this formulation also involves arbitrary choices, the largest of which is the selection of a relevant objective function. Other measures of relevance of the weights to a goal are possible and require reformulation of the solution methodology.

Table 6. June 1995 Lake Superior Outlook of Monthly Total Net Basin Supply (mm)[a].

Month	Quantiles 5%	20%	50%	80%	95%	Mean	Std. Dev.
Jun '95	99	108	149	167	188	141	30
Jul '95	80	101	114	142	166	120	26
Aug '95	44	82	95	131	151	102	35
Sep '95	-5	39	65	109	157	75	47
Oct '95	-5	23	46	77	93	49	30
Nov '95	-42	-14	2	30	66	10	33
Dec '95	-59	-39	-28	-15	2	-26	18
Jan '96	-65	-40	-23	-15	8	-25	20
Feb '96	-37	-22	-14	13	26	-6	23
Mar '96	-25	5	21	59	92	34	36
Apr '96	62	87	120	151	173	121	32
May '96	100	127	159	192	234	162	42

[a]Forecast non-exceedance quantiles, mean, and standard deviation, are expressed as over-lake depths. The quantiles, mean, and standard deviation are computed from the weights in Table 5. This hydrological outlook corresponds to the Climate Prediction Center *Climate Outlook* for June 1995, with probability settings on temperature for periods JAS, SON, DJF, JFM, FMA, and MAM, and on precipitation for JFM.

An important advantage associated with the computation of a weighted sample in the operational hydrology approach described herein is the independence of the weights and the hydrology models. After model simulations are made to build a set of possible future scenarios for analysis, several probabilistic outlooks can be generated with weights corresponding to the use of different climate outlooks, different methods of considering the climate outlooks, and alternate selections of just which of the 14 outlooks to use that are available each month. In making these alternate analyses and weights (re)computations, it is unnecessary to redo the model simulations to rebuild the set. This is a real savings when the model simulations are extensive, as is the case with Great Lakes hydrological outlooks. This also enables efficient consideration of other ways of using the weights to make probabilistic outlooks. For example, our use of non-parametric statistics in (5) restricts the range of any variable to that present in the historical record or in their hydrological transformations. An alternative that does not restrict range in this manner hypothesizes a distribution family (e.g., normal, log-normal, log-Pearson Type III) and estimates its moments by utilizing sample statistics defined analogously to those in (9). The detractor for parametric estimation is hypothesizing the family of distributions to use.

We built an *Outlook-Setup User-Product Interface* (or "*front end*" to our water resources forecasting system) as a specially-designed *Windows™* application. This allows the user to set hydrologic outlook parameters and to begin the hydrological outlook. The interface defines the hydrological outlook and historical-data periods, selects the periods, probabilities, and priorities of climate outlooks (newly available from the NWS Climate Prediction Center), and determines the method for considering the climate outlooks in making the hydrological outlook. NWS's climate outlooks can be particularly cumbersome and difficult to use; but this interface greatly clarifies and simplifies their use in making a hydrological outlook. It allows readily understandable user interpretation of climate outlooks and easy user assignment of relevant priorities.

In fact, the interface is so successful in allowing a lay-person to utilize NWS's *Climate Outlook* that we similarly built a derivative product (also a *Windows™* application) to allow anyone to directly use the *Climate Outlook* in their own applications. This interface makes all computations utilizing the new climate outlooks. It finds all necessary reference quantiles for using a climate outlook from a user-supplied file of historical daily air temperature and precipitation, sets up all climate outlook selections as boundary equations in (17), formulates the optimization problem of (18), and performs the sequential optimizations with either of two methods (either by using all historical data or by maximizing use of the climate outlook selections). The interface computer code is also available as a stand-alone FORTRAN implementation for use under a variety of operating systems.

7. ACKNOWLEDGMENTS

This paper is GLERL contribution no. 983.

8. REFERENCES

Croley, T. E., II, 1993: Probabilistic Great Lakes hydrology outlooks. *Water Resources Bulletin*, 29(5):741-753.

Croley, T. E., II, 1995: Using NOAA's new climate outlooks in operational hydrology. *Journal of Hydrologic Engineering*, (submitted).

Day, G. N., 1985: Extended streamflow forecasting using NWSRFS. *Journal of the Water Resources Planning and Management Division*, 111:157-170.

Hillier, F. S., and G. J. Lieberman, 1969: *Introduction to Operations Research.* Appendix 2: Classical Optimization Techniques, Holden-Day, 603-608.

Ingram, J. J., M. D. Hudlow, and D. L. Fread, 1995: Hydrometeorological coupling for Extended Streamflow Predictions. *Preprints of Conference on Hydrology*, 75th Annual Meeting of the American Meteorological Society, Dallas ,Texas, January 15-20, 186-191.

PREPRINTS

13TH CONFERENCE ON PROBABILITY AND STATISTICS IN THE ATMOSPHERIC SCIENCES

FEBRUARY 21-23, 1996

SAN FRANCISCO, CA

13TH CONFERENCE ON PROBABILITY AND STATISTICS IN THE ATMOSPHERIC SCIENCES

FEBRUARY 21-23, 1996 SAN FRANCISCO, CA

Sponsored by
AMERICAN METEOROLOGICAL SOCIETY

Front Cover: State of California annual precipitation anomalies for the period 1895-1994 as represented by Chernoff faces. Facial expressions graphically indicate the type of anomaly for a given year. Subjectively chosen were "frowns" for anomalous dry years, and "smiles" for anomalous wet years. The figure was created by Timothy J. Brown, Desert Research Institute, Reno, Nevada.

AMERICAN METEOROLOGICAL SOCIETY
45 BEACON STREET, BOSTON, MASSACHUSETTS USA 02108-3693